5 Genealogy as a Hobby

James W. Cortada

Many Americans have discovered that working out a genealogy table can be as much fun as solving a crossword puzzle.

—Peter Andrews (Andrews 1982, 17)

Introduction

The most widely read book in nineteenth-century America was "the family Bible." It was not called just "the Bible," but rather, "the *family* Bible." Between the seventeenth century and the 1960s, nearly every Christian resident of the United States became familiar with the Bible. It provided the only comprehensive set of perspectives on religious instruction, life, and worldview for many native-born Christian Americans. Much of the Old and New Testaments were genealogical, about generations of people, beginning (literally) with the story of Adam and Eve on the first page of the Book of Genesis.

Most Americans living in the nineteenth and first half of the twentieth centuries would have been familiar with passages in the Gospels of Matthew and Luke, which record the genealogy of Jesus Christ, and of his parents Joseph and Mary. Family linkages were discussed either through the daily routine reading of the Bible or in other forums. Bibles published in the United States even included several blank or preprinted pages placed between the Old and New Testaments on which families routinely recorded births, weddings, and deaths of relatives, often replacing their Bibles when these pages no longer had room for more names. It was not uncommon for a family Bible to be used for over a century to record such family events, and in some families, the practice continues. For those curious about their family history, these Bibles often provided a convenient starting point for identifying ancestors. Indeed, not until the early years of the twentieth century did state and federal agencies begin replacing this form of record

keeping with forms, most notably birth and death certificates, as true evidence of familial standing, necessary for documenting proof of identity, age, and details needed for passport applications.

When the U.S. Constitution went into effect in 1787, it mandated that a census be conducted every ten years by the federal government in order to apportion representatives to the lower house of Congress. Officials have conducted this census every decade since 1790. The census is conducted for political and operational reasons, not out of some genealogical consideration, but in the process the government has amassed a treasure trove of reliable genealogical information for over two centuries. These data are unique in the world and their value as a source of family information has increased over time.

In addition to counting people, from the very first census the government collected information about families: names of the heads of households, ages of all members of a household, and their gender. In subsequent censuses, enumerators collected additional information tied to families: names of children, occupations of all members of a household, relationships (parents versus children), and place of birth of parents. By the late nineteenth century, even American passport applications required the applicant's parents' names and ages, the applicant's place of birth, and both the applicant and parents' community of residence. Family and family history were tied together, even though the amount of data collected remained limited.

The family as an institution in society is part of one's worldview, a component of each individual's self-identity. This chapter explores how that awareness and, therefore, interest in one's family's past manifested itself in what today is often acknowledged as the nation's most widely indulged hobby—genealogy. But first some distinctions are needed. Genealogy is largely accepted to be the identification of one's ancestors: their names, their lifetime dates, their specific generational and familial relationships to living members of a family, and often where they lived and how they made their living. Family history is often intertwined with the concept of genealogy in that it too seeks to document previous members of a family, but also includes stories and evidence that describe the collective biography and handed-down values of all who came before us as part of the "family tree." Often the two terms are used interchangeably. The word hobbyist routinely is used to describe one who pursues genealogy, in part because so few of the many millions of people who engage in it have had any academic training in the subject. That fact is ironic, however, because genealogy is also a body of highly defined, formal best practices developed

over many hundreds of years into a disciplined profession and line of intellectual activity. The amateurs—hobbyists—and the professionals, and the discipline they share have had increased interaction from the start of the Gilded Age in post-Civil War America to the present, and are now also joined by some of the most sophisticated uses of information technology available today. How did all of that happen?

Who Are These People?

There is a growing body of evidence suggesting that genealogy has become an important pastime engaged in by many people. A few statistics suggest the extent of that engagement. As of late 2008, nearly three million people, the majority from the United States, had become active users of Ancestry.com, one of the largest Internet sites for genealogical research. Over nine hundred thousand have subscribed to its extended services, paying a monthly fee.

The number of Americans expressing interest in genealogy nearly doubled from the late 1970s to 2008. Interest surged when genealogical information became available through the Internet, with interest expanding by over a third between 1995 and 2001 (Bernstein 2001). These were not passive players. The Church of Jesus Christ of Latter-day Saints, which operates familysearch.org, one of the largest and most active online genealogical sites, reported that in the early 2000s it received in excess of eight million inquiries per day accessing its database of over one billion names (Kilborn 2001).

The profile of the person who engages in genealogy has changed over the decades. Formal study of one's family history dates to the early 1800s in North America with the formation of genealogical societies. These societies were located largely in New England, but subsequently spread, over the century, to the Mississippi River, and by the early 1900s across the continent. Early interest in the subject grew out of normal curiosity about one's family history, but also as a manifestation of social status through membership in patriotic societies that required documentation of one's genealogy. Wealthy and upper-middle class Americans of the nineteenth and twentieth centuries hunted distinguished relatives and even royal relations. Organizations that played important roles in encouraging such research, patriotism, and historical preservation included the New England Historical Genealogical Society (founded 1845), Daughters of the American Revolution (founded 1890), Colonial Dames of America (founded 1891), the General Society of Mayflower Descendants (founded 1897), the

National Genealogical Society (founded 1903), and the American Society of Genealogists (founded 1940 for professional genealogists only).

Throughout the twentieth century, the white middle class in the United States joined the wealthy and status-conscious in pursuit of their ancestors. They were curious about their origins and increasingly able to devote time and funds required to conduct such research. The amount of genealogical data increased during the last third of the twentieth century and was easily accessed in libraries and on the Internet. Ethnic groups, most notably African Americans, Hispanics, Jewish Americans, and Native Americans, then became involved. By late 2008, the number of genealogical sites devoted to ethnic groups had expanded rapidly: over seven hundred devoted to Hispanics and African Americans, some five hundred to Jewish families, and even some six hundred concentrating on Native Americans. In short, by the end of the twentieth century all major social, economic, and ethnic cohorts were actively participating in genealogical research.

Genealogy is considered a hobby. The price of admission is interest, time, and activity, much as with any hobby. Because one is not required to be licensed or trained (credentialed) to participate, there are many problems with factual errors in what data are collected. To be sure, formal training and resources are available—and there are professionally trained (even licensed) genealogists who conduct research for free or for fees, particularly in other countries where Americans need help tracking down information without having to travel overseas.

There are the inevitable questions about why genealogy is attractive to Americans. Answers stem from a variety of endogenous and exogenous forces at work. Personal reasons fall into four essential categories. First, there is the normal curiosity about one's heritage. "Where did I come from?" "Who were my people?" "Am I related to any famous individuals in history, or to some horse thief?"

Second, genealogical research has served as an excellent way to meet other members of one's family and to maintain connections with them. Genealogy provides a way to connect with relatives given the history of social and physical mobility that has characterized ever-migrating American families. The Internet's social networking and e-mail tools have made this even more possible in recent years.

Third, as individuals achieve a certain age they begin to think about their legacy and relations, about the past and future of their families. Increasingly, we read that Baby Boomers are particularly interested in genealogy, but the historical record suggests that in all decades elder

members of a family were often the most interested in passing on stories of their families to younger members.

A fourth motivation to engage in genealogy concerns the inspiration that emerges from family incidents and stories: the grandfather who fought at Normandy in 1944 telling his family about his Italian roots first in New York and earlier in Italy; African Americans attempting to identify their ancestral home in Africa; or others with tales of adventures in Europe, Latin America, or elsewhere in the world all serve as grist for the hobby. The discovery of documentation in an elderly relative's home is frequently an additional spur to investigation. One survey done in 2005 reported that 65 percent of those doing genealogical research became interested because of stories heard within their families, a percentage that seems low given the much higher percentage of people (85 percent) who have reported talking to relatives as their initial source of information (Market Strategies, Inc. 2005).

Exogenous forces, too, have long stimulated interest in genealogy and influenced how Americans approached the subject. There are three sets of such forces. The first comprise activities and influences immediately outside the family. Genealogical activity is largely centered within a family. Therefore, it is reasonable to consider influences outside the family as exogenous in order to understand this activity as a hobby. An essential collection of such forces are events that stimulate interest among those family members who conduct genealogical research, such as the anniversary of a national historic event.

A second category of exogenous forces influencing genealogical activity are various actions taken by governments, libraries, and other institutions either to make available or to deny access to information considered essential (or useful) to someone conducting genealogical research. The availability of a new decade's worth of census data in the United States is an example of this kind of force, as is the continued refusal of a government agency to make available certain types of data, such as personnel records for living relatives or tax returns.

A third force—and one that will receive more attention in this chapter than the first two—is the innovation in information technology (IT) that has profoundly affected genealogical activities in the United States, to such an extent that IT has become the most influential of the three exogenous forces.

Within the category of the first class of exogenous forces, the centennials of the American Revolution in 1876 and the Civil War from 1861 to 1865 led individuals to explore the roles of their families in those wars. In the years following the Revolutionary War centennial, genealogical and

historical societies came into being that served as major centers of research about family history. The Civil War's centennial in the 1960s led to a sharp increase in research concerning military records of soldiers fighting on both sides; the bicentennial of the American Revolution in the 1970s did the same. Both world wars of the twentieth century also stimulated interest in family history, although evidence is anecdotal.

A more direct influence was publication of Alex Haley's 1976 novel, *Roots*, which subsequently became a television series watched by some of the largest TV audiences up to that time (over 130 million viewers). One archivist at the National Archives reported increases in genealogical research of some 300 to 400 percent in its various centers following broadcast of the series (Haley 1976). The novel traced the history of an African American family, generating enormous interest on the part of people of all races and ethnic backgrounds to explore their ancestry, particularly African Americans, and often for the first time. Librarians all over the United States in the 1970s reported a substantial increase in interest in family history across all ethnic and racial lines, not just among African Americans, so much so that the *Roots* broadcast on television is widely considered by genealogists, historians, and librarians as a turning point in the history of American genealogy.

Where interest had long existed in genealogy, *Roots* did not necessarily encourage dramatically new efforts. For example, in the South, where interest in familial history and in the role of ancestors in the Civil War was and is widespread, *Roots* had less impact among white residents. However, it did have an enormous effect on African Americans. In short, each national event had varying influences on people doing genealogical research, depending on their ethnic background, local culture, and social status, suggesting that one must be cautious about broad generalizations of effects on amateur and professional genealogists alike.

Various events in American life continue to supply reasons for pursuing genealogical research: religious requirements (Mormons), patriotism (historical anniversaries), difficult times (world wars, the Great Depression), media events, and the possibility of conducting research in new ways. Exogenous forces of the second type—availability of new sources of information—continue to stimulate interest independent of some national event such as an anniversary or publication of a book. In the nineteenth century these included establishment of genealogical research centers and publication of the first "how to" books on the subject, followed by the creation of massive research collections, such as those of the Library of Congress (over 140,000 publications), the creation of online collections,

and most recently the study of DNA. Sometimes the surge in interest became substantial. For example, as a byproduct of the *Roots* phenomenon and in the same decade as the bicentennial of the American Revolution, the Library of Congress reported an 80 percent increase in the use of its genealogical resources between 1972 and 1977, largely by first-time researchers (New York Times 1977).

In 2001, twenty-five million records of immigrants and travelers entering the United States through Ellis Island at New York City between 1892 and 1924 became available online through the American Family Immigration History Center's Web site (http://www.EllisIsland.org). The site became an instant success, with well over eight million visits on the first day of its availability. The center possesses documentation on the names of people arriving in the country, as well as their countries of origin and when they arrived. Even the names of the ships that transported them have been preserved.

In 2008, the American government began making available a few income tax records from the early years of the twentieth century, which in time will probably lead to another surge of interest in family history.

The influence of newly available information in sparking research grew sharply once online genealogical services appeared, because they could rapidly spread news about additional accessible records to targeted audiences certain to want these kinds of announcements. Prior to such portals and services on the Internet, news about the availability of additional records spread more slowly, largely through word of mouth, when mentioned in a how-to manual, or when a researcher showed up at a library or archive that happened to have these materials. In the arcane language of the business manager, use of the Internet reduced the "mean time" to news and hence, the "cycle time" to process new data. In other words, researchers worked faster with the most currently available sources.

The Information that People Seek in Genealogy

What questions about genealogy do researchers seek to answer? Most are interested in knowing by name their blood relatives, going back in time as far as possible. They are keen to examine documentation that lists names of parents (e.g., birth certificates, passport applications, census records) and who married whom (e.g., marriage certificates and census records). They want to know birth, marriage, and death dates for each person and how relatives are related to each other. There has long been interest in military records of ancestors and, to a lesser extent, in wills and land records to

establish an ancestor's wealth and profession. Ties to churches, social organizations, and roles in communities also prove interesting. Of particular interest—and concern—is the changing spelling of a family's name over the centuries, which can either link families together or demonstrate that they are not related.

Once curiosity about direct lineage is satisfied, many individuals researching their family history become interested in acquiring stories of families associated through marriage. After this basic information is accumulated, researchers typically want to begin formulating a narrative of their family's past, such as what their ancestors did in Europe or Africa before coming to the New World and their subsequent roles after immigration to the United States. Weaving a personal history of a family into the much larger tale of the nation's past becomes the ultimate objective.

The epigraph at the start of this chapter strikes at the heart of the search for information, in that it can be quite varied, difficult to accomplish, and dispersed. Table 5.1 lists many types of widely consulted records essential to the genealogist. This list is by no means complete. Yet, note how varied and how many sources of information are available. Repositories of genealogical data can be voluminous, yet there might be few citations or even

Table 5.1

Variety of records consulted by genealogists

Birth certificates	Biographies
Adoption papers	Obituaries
Marriage certificates	Census reports
Divorce records	Military records
Death certificates	Church sacramental
Telephone directories	Association membership roles
School and university grades	Criminal records
Wills	Civil law suits
Family letters	Diaries
Emigration	Passport applications
Land deeds	Tax statements
Lineage organization files	Medical records
Newspapers	News and trade magazines
Oral histories	Professional
Photographs	Ship passenger lists
Tombstones	Voter registrations
Probate records	Family trees
Local history books	Histories of major events, battles

only one citation for a specific ancestor in literally thousands of pages of records. Genealogical inquiry is normally a tedious process that benefits from the use of finding aids, such as indices, and, most recently, software tools and the Internet to mine for specific data. Table 5.2 lists examples of the size and scope of the records various organizations hold that are essential for genealogical research, suggesting the mountain of material one might encounter in conducting such research. Table 5.3 presents a list of major Web sites accessed routinely by researchers today while table 5.4 lists widely used software tools.

Table 5.2

Size of genealogical archive, library, and online holdings, circa early 2000s

Library of Congress (LC)	40,000 genealogies, 100,000 local histories
National Archives and Records Administration	4 billion pages
Church of Jesus Christ of Latter-day Saints (LDS)	3.2 million microfilm rolls and microfiche; 356,000 books, serials, and other formats; over 4,500 periodicals; 3,725 electronic resources
Ancestry.com	1 billion names online
Cyndi's List	260,000 links to family history

Table 5.3

Widely used genealogical Web sites, circa early 2000s

Cyndi's List (http://www.cyndislist.com)

FamilySearch (LDS) (http:/www.familysearch.org)

Ancestry.com (http://www.ancestry.com)

Ellis Island (http://www.EllisIsland.org)

Genealogy.com (http://www.genealogy.com)

RootsWeb.com (http://www.rootsweb.ancestry.com)

Table 5.4

Widely used software tools

Family Roots (popular in 1970s to 1980s)

Ancestral Quest (early genealogical tool, still available in early 2000s)

Personal Ancestral File (PAF) (developed by LDS)

Family Origins (available in 2000s)

Family Treemaker (links to email)

Ultimate Family Tree

As of the late 1990s, *Family Tree Maker* had become the bestselling software tool of the decade, operating on PCs using Microsoft systems; it has proved relatively easy to use (Roberts 1998). Furthermore, all major libraries and archives also have Web sites and are increasingly making available online reference guides to their collections and providing access to key documents. One could just as easily have also cited the thousands of user manuals, guides, and articles published over the past two centuries that have been invaluable to researchers.

In most instances, the results from research that people need and want include five core categories of documents. These categories of information have remained remarkably consistent for over a century. The first category includes family trees, because they illustrate relationships among people within a family over time, usually including birth and death dates. These records, the minimum that researchers want, often prove so difficult to determine accurately that the effort to document them can consume a lifetime or more of work. A second set of desirable records consists of photographs of family members going back as far into the nineteenth century as possible; or prints, engravings, and portrait paintings, which can extend the visual record even further back in time. A third collection consists of myriad family records (e.g., legal and religious documents such as birth certificates, wills, and family letters). Fourth, equally prized, are collections of military records. These are well preserved in the United States, from the Revolutionary War to the present. Fifth, recordings of family voices, motion pictures, and videos of familial events and people, dating back to as early as the 1930s in many cases, have long been another source of documentation about the family. Families have been willing to spend considerable time and money to preserve these memories by moving film to video and most recently to DVDs, and transferring recordings from reel-to-reel tape to diskettes, then to CDs and DVDs. Businesses exist specifically to move this kind of content from one medium to another and to capture images off old photographs and slides and convert them into digital formats. Table 5.5 lists some examples of these businesses.

The most advanced form of family documentation is a written narrative of a family's history from its origins to the present. These narratives provide biographies of key family members, family stories and myths, handed-down practices within the clan, family members' roles in great and small events of a nation's past, and their trials and successes over time. Some family histories are informal: typed or even handwritten manuscripts that circulate within the family. Others are privately printed, and still

Table 5.5

Examples of U.S. media conversion firms, circa early 2000s

DigMyPics (http://www.digmypics.com) [largely photo and film scanning, conversion]

FotoBridge (http://www.fotobridge.com) [largely photo and film scanning, conversion]

ScanCafe (http://www.scancafe.com) [largely photo restorations]

QA Video Services (http://www.quavideoservices.com) [specializes in preserving home movies]

Digital Transfer Services (http://www.dtsav.com) [preserves video, film, recordings, and still shots]

Source: Online directory of digital restoration and preservation services, http://www .dmoz.org/Shopping/Entertainment/Recordings/Media_Preservation_and_Transfer/ Video/ (last accessed March 15, 2009).

others are preserved in the form of fictionalized novels and history books published by established publishing houses, such as Alex Haley's *Roots* by Doubleday and various academic studies of plantation families in the pre– Civil War South, published by university presses. The Library of Congress, long a collector of family histories, houses more than forty thousand such narratives in its collection. The ultimate goal of this kind of documentation is to tell "the story" of a family.

How People Conducted Genealogical Research before Personal Computers

While issues that genealogists and family historians investigate have remained essentially the same over time, the arrival of digital tools, beginning in the 1980s and assisted by the availability of personal computers (PCs) and early online files, began to change the way researchers worked to such an extent that one can reasonably divide the modern history of genealogical research into two periods, pre- and post-digital. However, this chronological division still requires us to acknowledge that techniques from the pre-digital era are still used today—such as discussions with family members and research visits to European churches—in part because much information remains undigitized.

In the United States then—and now—research begins with family members telling each other stories about the activities and life experiences of their own, then of their parents, and then of prior generations. This storytelling remains both the initial source of information about a family's past and often the key source of early inspiration for a member to pursue

genealogical research. Research often proceeds to books about a community's history, which are typically maintained in local libraries. Trips to ever-larger city and state libraries that hold larger collections of genealogical records, books, and periodicals, became standard practice by the late 1800s, and continued throughout the twentieth century.

Beginning in the 1930s, primarily federal institutions, but others as well, began to microfilm records such as census materials, making them widely available in regional libraries and at large universities. The widespread microfilming of newspapers, undertaken by the U.S. Library of Congress and many state university libraries, dating from the eighteenth century to the present, provided researchers with a new source. In fact, consulting census and newspaper microfilm (later microfiche) became two core activities of genealogical researchers. Those investigators who had the financial means might also hire a professional genealogist to conduct research, although extant evidence suggests that this practice remained quite limited (New York Times 1894). Thirty-four state genealogical societies that people consulted existed before Haley published *Roots*; eight more were formed afterward. But the number of professional genealogists remained low relative to the whole population of people doing genealogical research; indeed, most members of genealogical societies were hobbyists. To be sure, many were as highly skilled in the ways of genealogy as credentialed researchers.

Finally, one should acknowledge an old American practice, dating back to the eighteenth century: the how-to book for all manner of information and all kinds of work and play. Beginning in the nineteenth century and extending to the present, numerous books and articles appeared that guided researchers through various genealogical issues and sources of data, and advised on how to collect and organize information. Researchers in North America routinely turned to such publications to teach themselves the process, while classes and seminars sponsored by universities and local libraries provided a relatively minor source of guidance. An entire book could be written about genealogical guides, but suffice it to note that they had common features:

• Descriptions of what kind of information constituted genealogical data
• Lists of physical repositories of such information
• Issues related to accuracy of information from various sources, such as members of one's family
• How to create and manage a research agenda
• How to organize physical records and information
• How to write/tell a family's story

• Extensive lists of publications, libraries, and archives
• Discussions of special challenges, sources, and issues of specific ethnic groups, such as Catholics, African Americans, and Jews

These manuals often provided the only training researchers had unless, of course, they were professional genealogists.

Collecting information in pre-computer times proved to be a slow, tedious, time-consuming exercise. A piece of data would be picked up at a church or courthouse, off a tombstone, or in a family Bible or genealogical publication only to become an isolated fact that had to be matched to others so as to extend the narrative. A researcher's trips to various libraries to read microfilm consumed many weekend afternoons, while the volume of correspondence with librarians, archivists, and family members could prove massive. In fact, such correspondence was so great for so many researchers that authors of how-to books routinely devoted attention to its organization and use as part of the research process. In addition, a researcher's correspondence often became the occasion to exchange with family members or other interested parties records such as photographs, letters written by ancestors, and wills and other legal documents. Nothing was more exciting to a researcher than the arrival of a fat brown envelope in the mail containing several photographs of early forebears or a stack of old documents. Such a correspondence could go on for decades.

At the risk of oversimplifying the research process before the advent of computers, (which usually took hundreds of pages to describe in a user manual, varying by family and by issues pursued), a summary of what took place suggests how all these various tools were and continue to be used, serving as useful context for understanding the profound influence that information technologies have had on such activities. Initial research often started with relatives describing ancestors. That dialogue normally continued for the life of the project, usually for months, but often over many years and decades. Novices focused on identifying basic data, such as names of grandparents. More advanced researchers pursued answers to very specific questions that clarify existing information. There was always the race to query the oldest relatives before their insights and memories were lost. The next stage was often to acquire a user manual, study it, and then collect or consult the kinds of records already discussed (for instance, census or military records). Documentation proceeded to family trees, then added profiles of individuals with their birth and death dates, followed by details on individuals, such as their careers, civic activities, personalities, where they lived, and so forth. As progress was made, record keeping became important; that is when individuals begin creating voluminous

records stored in file cabinets, collecting material into three-ring binders, and writing to libraries and archives to build and verify the family narrative. Simultaneously, genealogists communicated with relatives to collect specific pieces of information and copies of documents, and to solicit help in visiting nearby libraries, archives, and courthouses to collect more data.

The process was highly iterative as one moved back in time, beginning normally with one's parents, moving next to investigating one's grandparents, then great-grandparents, and so forth. Along the way decisions were made about how broad to make the search (e.g., whether to build trees of aunts and uncles and their descendants) and whether to explore the roots of families that married into the family of the genealogist. New facts necessitated retracing steps to fill in gaps—just as one does in filling out a crossword puzzle—and to validate data now of questionable accuracy or to answer new questions.

This routine remained the unbroken practice from at least the mid-nineteenth century until information technologies began to emerge. While the IT tools described in the section that follows have affected profoundly how research is done and what kind of data can be collected, the core tasks and purposes remain unchanged. One still needs family trees, names, and relationships identified, birth and death dates, information about family events, and so forth. In the age of the Internet word of mouth remains as profoundly important as before, because much tacit knowledge lies locked only in the minds of elderly relatives, only to be revealed when asked the right question. People have traveled to libraries and archives all over the world from the 1800s on into the early 2000s, because the Internet has not eliminated distance from the life of a committed genealogist and family historian.

Research after the Arrival of Computers and the Internet

The emergence of information technologies made possible several fundamental changes in genealogical practices that represent the third category of exogenous forces at work. These changes affect research methods rather than the nature of issues addressed by researchers. A number of specific technological changes made computing accessible and affordable to researchers. Beginning largely in the 1960s, computer data storage expanded in size and dropped in cost by about 20 percent compounded each year, with the result that by the early 1980s, storing data on a computer was inexpensive enough for individuals to afford. That long-term trend continued to the present. One megabyte of memory on a computer in 1979

cost about $110,000; today it is less than $25. This evolution made storing massive quantities of information in computers a possibility for the even the most casual researcher.

A second development, beginning in the late 1970s but not widespread until the mid-1980s, was the repackaging of computing power into ever smaller and more affordable devices. The multimillion-dollar mainframe computers of the 1950s and 1960s made way for PCs in the 1970s and 1980s that could be acquired initially for up to $5,000 each; then came the $600 computers and laptops of the early 2000s. By 2000, PCs and other computer devices were small enough to be perched on a table or carried into an archive. They were also relatively inexpensive and many more individuals could afford to use computers to collect, organize, and display data in an effective and speedy manner.

A third and crucial technological development was the emergence of telecommunications linked to computers. Beginning in the 1960s commercial organizations began sending data over telephone lines from one computer to another, from terminals to mainframes and minicomputers. By the early 1980s, one could receive or send information from a personal computer to another PC, or access information from a large mainframe. When that function became reliable and practical, genealogists could conduct research through networks tapping into information in other computers without having to travel. At first they used dial-up telephone lines (in the 1970s and 1980s), then subscribed to consumer-oriented private networks (in the1980s and early 1990s), and then accessed the Internet directly through telephone or cable communications (beginning in the mid-1990s).

We have already alluded to the emergence of software designed to help genealogists. Three functions proved essential. First, software could be used to scan multiple sources of data available on various computers. This functionality was very limited in the 1970s and 1980s, for two reasons: technological limitations of the software and limited availability of information in computers. Google and other search engines did not become widely accessible until the mid-1990s, but nonetheless, searching was an early application of the software.

Second, software made it possible to organize information into databases (files, spreadsheets), beginning in the early 1980s on PCs and using software algorithms and methods developed originally for large mainframe files as far back as the 1960s. One could now store thousands of pieces of information without filling up rooms with paper, using methods that were effective in accessing this information.

A third development in genealogical software involved the display of data. Before computers, family trees were often messy, hand drawn paper representations, plagued by erasures and bad handwriting. In contrast, even the early software created family trees in neatly printed formats that could either be populated by hand, with people typing in the information, or eventually by using software to pull in data from online files stored in one's PC or available through such service Web sites as Ancestry.com. Over time these software-created trees became more detailed, moving from just names to names and dates of birth and death to each name being, in effect, linked to an electronic folder containing whatever information, pictures, and scanned documents related to that individual was available.

As with commercial computing in companies and government agencies, often technologies evolved faster than the content that went into them. However, beginning in the 1970s and extending to the present, the amount of information that governments, libraries, and researchers converted from paper and microfilm to digital formats became nothing less than a data tsunami. Key records in the United States, such as census materials, and other economic data moved to the Internet; churches set up Web sites and began populating these with birth, marriage, death, and burial records, as did state and county offices. The Church of Jesus Christ of Latter-day Saints began posting hundreds of millions of records in online formats. Hardly a week goes by without an announcement by some agency or organization in the United States (or Europe) that it has now made available new records over the Internet. The volume of electronically available material is so massive that it is nearly impossible to estimate the quantity. Attempts to do so have led many observers to provide statistical sound bites about how many online files now exist. That availability of online data has transformed most genealogical how-to books largely into IT user manuals, focusing on researching over the Internet.

By the end of the 1990s, therefore, Americans doing research had available the necessary IT tools and increasing amounts of digitized data they could use to construct family histories, along with essential infrastructure. In 1990 approximately 20 percent of American households had a PC; by 2000 deployment had exceeded 40 percent (and these PCs were connected to the Internet). As of early 2009, over 60 percent were Internet linked. However, if one takes into account the percentage of people online daily from any source—home, work, mobile device—then the statistics are higher: 50 percent in 2000, and nearly 80 percent in late 2008 (Pew Foundation 2009).

Simultaneously, the speed with which researchers could access information increased as they moved from slow dial-up lines to high-speed broadband in the early 2000s. At the start of 2009, well over half of the American public had access to broadband computing. Broadband connections made quick transmission of large volumes of information orders of magnitude greater than that which could be transmitted over a dial-up line. For example, in the 1990s one could download a photograph of a relative to one's PC over a dial-up line in about five to six minutes. In the early 2000s, the same photograph, along with several dozen others, could be transmitted over a broadband connection in less than half a minute. The convenience and ease with which this important activity could take place encouraged more such transactions, reducing the need to wait for the U.S. Postal Service to deliver those brown envelopes so anticipated in the past. Table 5.6 lists major uses of the Internet for genealogical research as of the early 2000s. Table 5.7 shows types of information that became available over the Internet since approximately 2003, reflecting the current pattern of ever-more varied materials becoming available.

As information became more accessible, behavior by researchers began to shift, moving from activities of lone genealogists to more family-centered projects. To be sure, in the 1980s and early 1990s, individual researchers worked in the same ways they had before, simply grafting use of PCs onto their existing practices. As email usage spread in the early 1990s, the amount of communication one could engage in with relatives increased, leading to expanded collaboration within families. The extent of that expanded dialogue has yet to be studied. Mounting evidence suggests

Table 5.6

Major uses of the Internet for genealogical research, circa 2003

Search indexes, vital records, and statistics of individuals (birth, death, marriage, census, immigration)

Search classified directories and collections of Web genealogical resources

Create personal family-history pages

Explore searchable databases of personal family-history pages, commercially collected data products, indexes of government-collected data sets

Participate in genealogical discussion groups

Use genealogical freeware and shareware

Access city, county, and state genealogical society Web sites

Access individual library Web sites and online catalogues

Source: Diane K. Kovacs, "Family Trees on the Web," *American Libraries* 34 (7) (August 2003): 44.

Table 5.7
New types of data available over the Internet, since 2003

Maps and other GIS data sets
Ship passenger lists
Cemetery transcriptions
Wills, deeds
Tax files
Newspaper runs from earlier decades
Genealogical journals
Personal family histories
Business directories published by cities and professional associations
Discussion groups

that a new activity grew in the 2000s: social networking. Web sites such as Ancestry.com and others that used social networking tools added MySpace-like community functions and bulletin boards that made it possible, for example, for multiple family members to populate simultaneously a family-wide tree database, working from various locations, updating the tree in real time. In recent years, social networking has facilitated the sharing of family trees, photographs, and other documents, as well as the creation of birthday calendars.

As Americans began to use commercially available Web sites, a Wiki-type of behavior emerged, recommended by almost all the Internet-oriented how-to manuals of the 1990s and 2000s. These provided the capability to create online files housed at such sites as Ancestry.com. Others could add content, correct errors, and make material available to other individuals. Tasks were performed by Web site users, not by site providers. These tools resulted in significant user adoption. For example, between July 2006 and October 2008, management at Ancestry.com recorded:

• More than 7.5 million family trees created
• More than 725 million profiles added to those family trees
• More than 6.5 million Ancestry.com users created family trees
• More than 10 million photos uploaded into Ancestry.com
• Nearly 150 million Ancestry Hints accepted (Bonner 2008)

These statistics are useful indicators of this Internet-facilitated behavior, because it occurred within the largest genealogical Web site in the United States. Rival services could reasonably be expected to report similar increases in activity, since the ability to post data at these sites is a core function

offered. It is a short step to yet another community-oriented action: creation of family Web sites.

Families increasingly set up their own Web sites, a practice that began around 2005, within existing social networking tools such as MySpace and more frequently, Facebook. Often these family sites were created to facilitate communications among relatives and to share photographs. Younger tech-savvy family members created these sites for such activities as weddings (e.g., www.eWedding.com, www.Wedsite.com) and other family projects; some then evolved into ongoing family Web sites. As Americans over the age of 45 (traditionally the age group most interested in genealogy) began interacting with these social networking based sites, they began transforming into genealogical Web sites. More members of a family could participate in a collaborative way in conducting genealogical research, learning more family history and even organizing family reunions in the process.

In the period following wide availability of Web browsers (post-1995), which made use of the Internet dramatically easier for most people, genealogical research led to an unanticipated consequence: people discovering lost friends and relatives. Both white and yellow page telephone directories in the United States became increasingly available online in the early 1990s. Over time, various local directories were interconnected so that by the end of the 1990s one could look up, for example, everyone in the United States whose last name was "Cortada," or everyone with the name of one's lost college roommate.

Services appeared that performed data mining searches, for instance, of high school classes by year of graduation, for a fee. It did not take long for family historians to use the same tools to find lost or new relatives. As one observer noted in 1999, with online genealogical tools, including telephone directories, "many who became separated from their families, and thus their family histories, have discovered that they can find themselves reunited with long-lost cousins and long-forgotten lore" (Howells 1999). Nearly a decade later, observers still reported expanded use of the Internet to thicken connections within families and to bring new participants into family discussions, so that the "digital age family history has become a sociable, co-operative movement" (Hudson and Barratt 2007, 20–21).

To summarize, research methods involving use of the Internet extended the range of pre-Internet work: digitization of records, extensive access to noncommercial and commercial Web sites, establishment of family social networking sites, extended communications among family members, friends, and other hobbyists for sharing and exchanging

research, collaboration among larger groups of people, discussion and information requests from a larger pool of potential informants, and socialization in general. The most immediate results that the emerging Internet use had on genealogical research can also be summarized. Most obvious is the increased speed with which information could be accumulated, using the power of software search engines and telecommunications. Mail delivery speeds became increasingly irrelevant. The amount of personal travel to libraries and archives declined as online researching reduced the need to physically visit sources of information. Internet users could prepare for a research trip far better than before, because so much preparatory information could be examined in advance, such as the online indices of records to be examined.

Yet other behavior did not change. For example, hobbyists still bought books on how to use the Internet, teaching themselves how to conduct genealogical research. A search of Barnesandnoble.com, the online bookstore, in 2009 indicated there were several thousand books for sale on or about genealogical matters, while Amazon.com listed over one hundred titles available on its e-book, Kindle. Emblematic of Americans' interest in relying on books and other instructional materials was Terri Stephens Lamb's book, *E-Genealogy: Finding Your Family Roots Online,* published in 2000, which addressed the reader in the first person and speaks of "easy-to-read steps" intended to "show everything you need to know to bring your heritage home," personalizing the obligation to learn the techniques through use of the manual (Lamb 2000). This book is a useful example of what had been learned in the previous half-dozen years about using Internet-based genealogical tools and is a good representation of practices that have continued a decade after publication.

The Internet Did Not Change Everything

As important as the Internet became in augmenting genealogical research, it has not proven to be the be-all and end-all for family historians. One major problem that defies resolution is the spread of misinformation. Beginning in the 1800s, experienced genealogists warned about sloppy research and emphasized the need to verify data, even if from official sources or family Bibles, because records are collected inaccurately, people lie, information is lost, and bad data survives and circulates perniciously. The Internet facilitates dissemination of false facts in much the same way as paper records.

Elizabeth Powell Crowe, author of one of the most widely respected guides about the use of the Internet in genealogical research, still felt compelled to comment extensively about trusted sources as late as 2002. She almost could have written the same way about published sources a century earlier: "Most serious genealogists who discuss online sources want to know if you can 'trust' what you find on the Internet. Many professional genealogists I know simply don't accept what's found on the Internet as proof of genealogy, period. Their attitude is this: A source isn't a primary source unless you've held the original document in your hand" (Crowe 2002, 12–13). Mimicking standard good practice of historians, she commented that even a paper source should not to be fully trusted unless its data are confirmed by another paper source. "In my opinion, you must evaluate what you find on the Internet, just as you evaluate what you find in a library, courthouse, or archive" (Ibid., 14).

Genealogical scams, such as companies offering to sell family history, certificates of authenticity, and family heraldry have been active for decades and continue to be a problem for novice researchers. So, normal discretion and a healthy skepticism remain essential best practices for a family historian—behavior that is infrequently evident. Thus, for example, the millions of contributions of content to Ancestry.com's databases are not all guaranteed to be accurate.

A second problem that persists is the existence of material not yet on the Internet, which means one still must travel to U.S. courthouses around the country, to church archives across Europe, and to libraries all over the world. Only a fraction of all relevant materials are online. For example, the Roman Catholic Church has kept records of who received the sacraments (birth, first communion, confirmation, marriage, and death) for nearly two thousand years, and made it a formal responsibility of parishes since the late Middle Ages. To be sure, wars, fires, sloppy management or workmanship by a semi-illiterate priest in one church or another compromised the process from time to time. Nonetheless, billions of records exist in Catholic parishes and libraries all over the world that have yet to move onto the Internet. Catholic Church records are absolutely an essential source of genealogical information for Americans probing their families prior to their arrival in the New World, and even for early migrants for hundreds of years in Latin America and the Caribbean. Table 5.8, therefore, is the flipside of table 5.7, showing what is not yet widely available on the Internet.

This leads to a related discussion, of what the Internet still cannot do. The Internet cannot be used to communicate directly with a source or

Table 5.8
Sample types of data not widely available over the Internet, circa 2008

Most Catholic Church records
Most local European newspapers
Most U.S. county court records, pre-1980
Post-1930 U.S. Census records
Vast majority of documents in state and national archives
Municipal records from Eastern Europe and Asia
Colonial records from European country archives

family member who does not use the Web or even anyone who does not use email. Finding information on many family names, which may have various spellings, remains tacit knowledge (that artificial intelligence has yet to grapple with). Some names cannot be found on the Internet; they have not migrated there. It is difficult to effectively verify the accuracy of information found on the Internet; that remains a very important human activity of discernment and critical judgment, informed richly by the context of known family history.

Nearly as frustrating is the fact that there is cost associated with placing accurate copies of original documents on the Internet, so relatively few accurate copies of original documents are available. Further, original documents will only be obtained via online application forms made available by reliable sites, such as online records of applications for passport originally submitted as an e-service by a government (such records are not available today).

Other aspects of genealogical research that the Internet only partially facilitates are the contextual conversations and visitations so essential to understanding a family's history. Talking to elderly residents of a village about life in the community in which one's ancestors lived can only be done by visiting the town and sitting down with the "old men in the park" to understand what it was like during a critical period, such as in war, famine, and economic boom times. For example, visiting a rural locality gives one an enormous sense about lifestyles there, including how far home was from work, school, and church; and the effects that weather and soil conditions had on a family's prosperity and worldview. Just as military historians universally insist that one must walk a battlefield to understand how a battle was fought, won, or lost, experienced family historians walk where ancestors trod to understand more deeply their prior experiences. On such trips one can photograph "ancestral homes," have a face-to-face

meeting with a distant relative who gives the genealogist the gift of some old papers and photographs, discover a school textbook used by an ancestor in the 1800s, and buy local histories not listed by online book dealers such as Amazon.com. We still live in a time when digital means are used only as part of a much larger set of tools to do genealogical research.

A New Frontier: Research after the Internet

In 1998, a group of African American descendants of slaves belonging to President Thomas Jefferson proved through DNA testing that they were his blood relatives, tracing their ancestry back to Jefferson's slave mistress Sally Hemings. This topic became the first major historical subject of debate that turned on the issue of DNA testing in genealogical research. The effect of this finding proved electrifying among African Americans interested in their genealogy, much as *Roots* had stimulated their interest in the 1970s. This discovery opened a new chapter in genealogical research on the possibility of using of DNA to link families together through biological proof made possible by advances in genetics research.

During the 1990s various groups took early steps to conduct genographic research, communicating largely through the Internet with interested researchers. However, it was not until the early 2000s—when a world-wide study was initiated to determine the DNA relationship of all humans and, more specifically, to prove that all people were related and had come out of Africa sixty thousand years ago—that general interest in the subject expanded (Behar et al. 2007). IBM and the National Geographic Society collaborated with Spencer Wells, a geneticist and anthropologist at National Geographic, to collect specimens of DNA from various groups of humans all over the world. To offset partially the project costs, National Geographic publicly invited everyone worldwide to donate a small fee and their DNA to the effort and receive in return a summary of migratory patterns of other and earlier people who shared their types of DNA.

The project was accompanied by television publicity, press coverage, and extensive discussion over the Internet; hundreds of thousands of individuals participated almost immediately. Various other organizations began offering similar services. For example, Genetic Genealogy, DNA Tribes, and Familybuilder all merchandize their services primarily over the Internet. At the moment, the premier project is National Geographic's. More to the point of our story, this is a new use of technology and science in support of genealogical research and, as with earlier computer tools, is quickly being embraced by family historians.

Conclusions

This book is devoted to a discussion of how Americans find the information to support their everyday activities. The experience of genealogists reflects a long-standing practice evident across many facets of American life, of incorporating each new form of information-handling technology that has appeared in support of work, war, play, religion, politics, sports, and hobbies. It is a theme that the author and others have argued is an important feature of American society in evidence since colonial times (Chandler and Cortada 2000). Internet users rely on the Web at the same time that they rely on all manner of preexisting tools and sources of information. Thus, it is not unusual for American genealogists to go to the Web for some sources of material, visit their bookstore for how-to manuals, and make pilgrimages to courthouses, archives, and libraries to look at books and actual documents, talk to experts face to face, and spend many hours reading microfilm and microfiche. Researchers take their notes on 8½ × 11-inch pads, 3 × 5-inch index cards, notebooks and three-ring binders as well as using laptops, personal computers, digital cameras, and photographic functions on their cell phones. They scan photographs, store jpg files of these pictures, and go to their neighborhood Walgreen's pharmacy to pick up paper copies of images emailed to them by distant relatives. One other piece of evidence in support of the notion that genealogists simultaneously use multiple technologies: The vast majority of how-to books on genealogy published from about 1995 to the present in the United States include either a CD or DVD with various software tools and forms. Not even technical manuals published for computer programmers were adorned as comprehensively as those designed for genealogists.

Americans value speed of execution, and it is one of the enduring features of the Internet that individuals can gather large amounts of information about their families in a very short period of time, accomplishing in hours what used to take years. This can be accomplished from the comfort of their home offices, using a laptop connected wirelessly to well-stocked Internet sites such as Ancestry.com or that of the Church of Jesus Christ of Latter-day Saints. Convenience increased activity, as the reports from commercial Web sites, archivists, and librarians suggest. Sales and contents of how-to manuals also reflect this growing use of the Internet by which individuals and families explore their past. One could argue that this speed of execution and do-it-yourself approach to genealogy across all social, economic, and racial groups reflects the larger American practice

of applying "sweat equity," similar to the way in which they personally remodel their homes or restore old automobiles and boats.

We live in a data-hungry world. At one time it was enough to have the family tree decorated with the names of ancestors, and the years of their birth and death. The rest of the information was largely stored in people's minds and passed down through oral traditions and narratives told by grandparents to grandchildren, occasionally aided by the fact that someone forgot to clean out a box or trunk in an attic containing papers and photographs. Even the most casual viewer of the popular public television program *Antiques Road Show* can be amazed at how much material is saved accidently by seemingly everyone. Serious genealogical researchers, indeed often the professional genealogists, began arguing in their professional publications, in presentations, and in the press, for more context and information about individuals, beginning in the late 1800s; not until the last third of the twentieth century did this imperative become possible for the masses.

One of the features of Internet-based genealogical research is the massive diversity of information people want to explore. Hobbyists want important collections put online. When the Internal Revenue Service announced in early 2009 that it would begin to make a few tax records available to the public largely from pre-World War I, reaction was less "wonderful news!"—although it was news—and more "how can we access it over the Internet?" The thought of having to go to some regional National Archives library to read microfilm now seemed positively primitive. As in other areas of their lives, Americans became accustomed to searching myriad databases for information at work and play, so they applied the same practice to their study of family history. The literature on how to conduct that kind of research and studies about genealogical practices in the United States, many cited in this chapter, leave behind a documented trail of evidence that clearly shows an expanding appetite for varied information that can inform family history. Indeed, the whole notion of moving from genealogical or family trees to family history is symptomatic of the change made possible by various types of information technology, including, but limited to the availability of the Internet.

Another feature of how genealogists use the technology—and one not discussed by observers—is the impact users of Internet sources are having on the technology itself. So many people are using the Internet to do genealogical research that enterprises serving these consumers can justify the cost of creating new software tools and augmenting search engines' mathematical underpinnings and capabilities, as evidenced by advances in

major genealogical Web sites. Economies of scale and the volume of demand provide markets for new tools, such as the three-dimensional family trees that include, within each name on a tree, paths to sub-data-bases that hold images of photographs and original documents, as well as video and audio files. Functions associated with family trees are far more advanced (as of 2009) than many of the data management tools and portals available in more traditional business, teaching, and government applications. The functionality provided has more in common with the data-intensive research projects one finds in pharmaceutical genomic research, the hunt for new medical compounds, or modeling of weather.

As soon as new functions become available to the genealogist they are rapidly adopted. New requests made for further advances. Like other multifaceted data-intensive applications, those of genealogists are pushing information technology forward across a broad band of activities: database management, data mining, data visualization, graphical representation, memory systems, computer cycle speeds, and, of course, demand for access to an ever-larger capacity of broadband communications across the entire country. The only other corner of American society where so much innovation is routinely possible exists within the U.S. Department of Defense, where, for example, requirements to train annually several million people, most of whom are under the age of thirty and accustomed to video games and screens, makes it affordable to do ground-breaking research and innovation in IT-based teaching and real-time automated high-performance systems, such as battle-field management. Other civilian applications described in this book are also contributing to the evolution of the Internet itself, and it is the confluence of multiple types of uses and transformations that affect genealogical work. One final point about the power of economies of scale: the monthly fee to use such advanced tools at a major genealogical site is as little as $12 to $30 dollars per month (as of 2009).

Since many of their work activities involve use of computers, it should be no surprise that users would bring over to genealogical activities the same tools and methods to other aspects of their lives including genealogy. For example, a genealogist might use the laptop assigned by their employer to access a genealogical Web site and perhaps write up search results using Microsoft Word that he or she learned to use at school. Organizing materials or conducting research in the manner of a management consulting project, as many how-to books suggest, grew out of practices in wide use in manufacturing, military, and research functions and industries. Given what we know about the demographics of those conducting genealogical research today (for example, employed Americans age forty-five and older)

we can assume that the majority learned to use the Internet not by doing genealogical research but by using computers at work. Over 80 percent of all workers interact with computers in one fashion or another, and we are rapidly reaching the point where over 90 percent of all people between the ages of ten and forty-five interact with the Internet.

Acknowledgments

This chapter was built upon the research of Michelle Bogart and Holly Green for which the author is deeply grateful.

Bibliography

Andrews, Peter. 1982. Genealogy, the Search for a Personal Past. *American Heritage* 33 (August–September): 17–21.

Behar, Doron M., et al. 2007. The Genographic Project Public Participation Mitochondrial DNA Database. *PLOS Genetics* 3 (6): 1083–1095.

Bernstein, Elizabeth. 2001. Genealogy Gone Haywire. *Wall Street Journal*, June 15, W-1.

Bonner, Susanne M. 2008. Email to author, October 21.

Chandler, Alfred D., Jr., and James W. Cortada, eds. 2000. *A Nation Transformed by Information: How Information Has Shaped the United States from Colonial Times to the Present*. New York: Oxford University Press.

Crowe, Elizabeth Powell. 2002. *Genealogy Online*. New York: Osborne/McGraw Hill.

Haley, Alex. 1976. *Roots: The Saga of An American Family*. New York: Doubleday.

Howells, Cyndi. 1999. Tracking Your Family through Time and Technology. *American Heritage* 50 (1). http://www.americanheritage.com (accessed May 1, 2010).

Hudson, Jules, and Nick Barratt. 2007. The Rise of Family History. *History Today* 57 (4): 20–21.

Kilborn, Peter. 2001. In Libraries and Cemeteries, Vacationing with Ancestors. *New York Times*, August 19, 1, 24.

Lamb, Terri Stephens. 2000. *E-Genealogy: Finding Your Family Roots Online*. Indianapolis, IN: Sams Publishing.

Market Strategies, Inc. 2005. Market Strategies Survey. Unpublished, in author's possession.

New York Times. 1894. Its Twenty-Fifth Birthday: New York's Genealogical and Biographical Society. *New York Times*, February 19.

New York Times. 1977. Tracing Their Roots. *New York Times*, February 20.

Pew Foundation. 2009. Pew Internet & American Life Project, Pew Research Center. http://www.pewinternet.org (accessed May 1, 2010).

Roberts, Ralph. 1998. *Genealogy via the Internet: Computerized Genealogy*. Alexander, NC: Alexander Books.